Under-explicit and minimally explicit reference: Evidence from a longitudinal case study

Jonathon Ryan

Waikato Institute of Technology

jonathon.ryan@wintec.ac.nz ORCID: 0000-0003-0980-6900

Introduction

Underpinning a great deal of communication are references to people, in which speakers use a referring expression (RE) to clarify which individuals they mean. In doing so, they must select an expression that is not only grammatically and semantically permissible but is also pragmatically appropriate for various aspects of the discourse context. This dual processing demand, lying as it does at the confluence between pragmatics and grammar, has proved an intriguing site for SLA research. However, despite the substantial body of previous research, to date there have been few longitudinal studies involving post-intermediate learners, and very few based on data elicited in ways other than prompted production tasks; the present study works within this research space, presenting a longitudinal case study of one Korean user of English.

Since definitions of reference can vary considerably, it is worth specifying exactly which phenomena are to be examined here. Following Bach (2008), a rather restrictive definition has been adopted, equating to the Conversation Analysis term *recognitional reference* (Sacks & Schegloff, 2007), in which the speaker intends for the addressee to identify the real-world referent. To be in a position to do so, the addressee must have prior knowledge of the individual; for present purposes, this includes knowledge through prior mention of an otherwise hearer-new individual. Consequently, reference is nearly always achieved through use of a definite noun phrase (NP) (Bach, p. 28). Not included under this definition are mentions of hypothetical individuals (e.g. *the next person you see*), generic NPs (e.g. *the first born in a family*), introductions of hearer-new individuals (e.g. *a doctor*). For present purposes, the focus is further restricted to singular third-person references.

In addition to this definition, the present study adopts Accessibility Theory (AT) (Ariel, 1990, 2001) as its key linguistic framework (see the introduction to this volume for an overview). The key idea behind the AT model is of a hierarchical arrangement of NP types that maps to a hierarchy of accessibility. In this way, zero anaphora (ϕ) and pronouns specialise as *high-accessibility markers* (HAMs), signalling that the referent is readily recoverable from short-term memory; conversely, *low-accessibility markers* (LAMs) such as proper names and *the* + noun NPs indicate that the referent must be recovered from long-term memory. Referent accessibility is determined by the sum of various weighted factors, principally the distance between anaphors and antecedents, the presence of other referents (competition), salience (e.g. topicality; physical presence) and unity (e.g. the effect of discourse boundaries) (Ariel, 1990, 2001).

Accessibility marking in L2 reference

A number of studies have reported findings relevant to the developmental trajectory of L2 reference, typically through cross-sectional studies comparing features of reference by learners at different language levels (e.g. Crosthwaite, 2013; Nakahama, 2009; Takeuchi, 2014). Several other studies have included a longitudinal analysis, including Lumley's (2013) study of English-speaking learners of Japanese (2 measurements, approximately 18 months apart), Kim's (2000) study of Korean learners of English (2 measurements, either 8.5 or 13 months apart, and Broeder (1991) and Klein and Perdue's (1992) studies of developmental patterns in low-level learners (3 measurements over 2 ½ years).

Although there are important and undoubted effects of cross-linguistic influence (see for example Jarvis, 2002; Nakahama, 2009, 2011), the weight of evidence suggests there are certain developmental patterns that occur largely irrespective of the configuration of source and target languages. This trajectory was depicted by Chini (2005) as beginning with an early 'pragmatic and lexical' stage, characterised chiefly by alternations between bare nouns or names, and zero anaphor (Ø). The latter is used where the referent is easily identifiable from context, and appears to chiefly occur for reference maintenance in topic position (Chini,

2005; Klein & Perdue, 1992). Also acquired early are deictic expressions. Thereafter, further forms are incorporated, beginning with (if they exist) anaphoric pronouns and marking for (in)definiteness. At intermediate levels, an '(over-) explicit lexical' stage is reached, whereby learners overuse full NPs in contexts in which pronouns or ø would be appropriate. This is followed by a more advanced 'syntactic' stage at which learners have greater control over syntactic devices for maintaining reference, such as passive voice (Nakahama, 2011) and (in some languages) clitics (Chini, 2005).

In terms of accessibility marking, this trajectory is from one of *under-explicitness* at beginner levels, to *over-explicitness* at intermediate and low-advanced levels (see also Ahrenholz, 2005; H.-Y. Kim, 2000). These terms relate to mismatches between the accessibility of the referent and the accessibility signalled by the RE. An expression is under-explicit when it indicates higher accessibility than warranted, such as when a pronoun is used where a name would be felicitous; it is over-explicit when it indicates a lower accessibility than warranted, such as the use of a name in place of a pronoun.

Under-explicitness, although most characteristic at lower levels, persists at higher levels mainly in relation to referent introductions (Nakahama, 2003; Ryan, 2016). To date, evidence from referent tracking relates mainly to the use of ø in place of pronouns (e.g. Lumley, 2013; Nakahama, 2011; Ryan, 2012). However, there are exceptions. In particular, Lozano (2009, 2016, 2018) reports occasional infelicitous use of ø to mark topic shift among advanced English and Greek learners of Spanish, while the author of the present study found one Chinese participant greatly over-using pronouns in place of full NPs (Ryan, 2015).

By contrast, over-explicitness characterises a great deal of reference at intermediate to lowadvanced levels. It is illustrated in the following extract, in which *the young lady* is used in place of *her*.

so the policeman um ran after the lady and ø caught the la- the young lady um at last, and ø send the young lady to the tr- truck. (Ryan, 2015, p. 847)

As shown in this example, over-explicitness tends to occur more often in focus position or when the referent shifts between focus and topic position. It is not adequately explained by cross-linguistic influence, since it occurs seemingly irrespective of the L1 and L2 in question, being reported for – among others – Chinese learners of German (Hendriks, 2003), Turkish and Moroccan learners of Dutch (Broeder, 1991), English learners of Japanese (Lumley, 2013), Dutch learners of French (Gullberg, 2006), English learners of French and French learners of English (Leclercq & Lenart, 2013), and Mandarin and Korean learners of English (Crosthwaite, 2014). It is also not accounted for by learners, creating fewer opportunities for the use of high-accessibility markers (Ryan, 2015). Among proposed explanations are the possibilities that choosing fuller NPs eases processing load (Chini, 2005; Gullberg, 2006); that it is a means of avoiding pronoun errors (Gullberg, 2006); and that it is a strategy to promote clarity (Leclercq & Lenart, 2013; Lumley, 2013; Ryan, 2015). The latter is in keeping with Lozano's (2016) recently proposed Pragmatic Principles Violation Hypothesis (PPVH), which holds that infelicities resulting in over-explicitness are weak violations of Gricean principles with limited communicative consequence, while under-explicitness represents a strong violation at risk of triggering miscommunication; when in doubt, learners will opt for the weaker violation.

Despite these general tendencies across learner populations, as with other areas of learner language, studies of L2 reference must remain alert to the possibilities of cross-linguistic influence. Such influences are evidenced in the production of errors and in the acquisition of new RE types (e.g. Crosthwaite, 2013; Nakahama, 2011) but to date are less clear in terms of the felicity of accessibility marking. For instance, even where there is an identical referential feature in the L1 and L2, such as use of ø in subject position in Greek and Spanish, learners may routinely use infelicitous forms until advanced levels (Lozano, 2018). More generally, the relationship between L1 endowment and L2 output proves complex and far from simply a matter of language transfer (e.g. Odlin, 2003).

In relation to the present case study, the relevant cross-linguistic comparison is between English and Koreanⁱ, where one relevant observation is the availability of (but highly infrequent use of) third-person pronouns in Korean. In reference maintenance contexts, where English speakers use pronouns, Korean speakers alternate between the use of ϕ and either

bare NPs (Crosthwaite, 2014; H.-Y. Kim, 2000) or full names (S.-H. Kim, 2013; Song, 2005). An unmediated transfer of this strategy to English would lead to (1) under-explicitness by way of infelicitous use of ø; (2) over-explicitness by way of infelicitous use of names and bare nouns; and (3), under-use of pronouns overall. These predictions are, in fact, generally in keeping with previous findings for L1 Korean L2 English (Crosthwaite, 2014; Kang, 2004; H.-Y. Kim, 2000) but are also generally true for intermediate/advanced L2 reference. However, as will become apparent, they are not borne out in the present findings.

The current study

This study aims to explore the (under-researched) area of development in accessibility marking from a longitudinal perspective. Although the overall developmental trajectory appears well-established, it is unclear whether, for example, pragmatic development in an ESL context is best characterised by steady growth, alternating periods of growth and stability, or even periods of backsliding, perhaps as new RE types and strategies are added to the speaker's repertoire. With such issues in mind, the following research questions were posed:

- What changes are evidenced in the participant's RE system over the period of the study?
- 2) How does the participant's accessibility marking evolve over this period?

The first question focuses particularly on the RE types that the participant used, including errors of form and the expansion or contraction of this repertoire. The second question focuses on the issue of pragmatic felicity, and in particular longitudinal evidence of over- and under-explicitness.

Methodology

Participant

This longitudinal case study focuses on interview data from one Korean learner of English, Yoona. Yoona was recruited as part of a wider study involving 12 participants and four researchers, aiming to track the experiences of international students in mainstream study.

At the time of the first interview, Yoona was 24 years old and on the second day of a bachelor degree program in New Zealand; the final interview was shortly prior to her final examination nearly 2¹/₂ years later. She had previously graduated in the same field in Korea, where she also had some work experience. Upon arriving in NZ, Yoona studied academic English programmes in two different institutions for a total of 13 months, successfully completing an English certificate programme with an exit level equivalent to IELTS 6.5, placing her at the threshold between an upper-intermediate and advanced level of English. Despite her writing and reading results being strong, in the first interview she spoke at length about her concerns with her level of spoken English. However, her mainstream studies involved very rich language input, including a large amount of academic reading, essays and other written assignments, regular lectures, group work, oral presentations, and also several workplace internships. She also virtually always spoke English on campus (though never at home with her Korean-born husband and family). Thus, by the end of her studies, she felt that her communication skills and general confidence in using English were greatly improved, and this accords with the author's general appraisal based on comparisons of the interview recordings.

Data elicitation

Yoona participated in 18 interviews with the researcher between July 2016 and October 2018, with these lasting generally for between 45 minutes and an hour. The interviews were very loosely structured and free flowing, framed around the general question of 'how are your studies going?' As such, although elicited, these are authentic interview data designed to discover aspects of Yoona's experiences rather than solely to elicit acts of reference.

For the purposes of the analysis of reference, the first, middle, and final interviews were selected for detailed transcription. As it happened, the middle recording (August, 2017) is the shortest of the 18, being only 22 minutes long due to an interruption. To supplement it, the first 15 minutes of the following interview (September, 2017) were also included. These 15 minutes captured a full retelling of a series of events involving classmates and tutors, and is thus rich in references. Extracts from the three time periods (July, 2016; Aug-Sept. 2017; Oct. 2018) include roughly similar numbers of referent introductions and acts of referent tracking.

Analysis: Accessibility coding

In the first stage of analysis, third-person singular references to people were coded as either 'introducing' or 'tracking', and further coded by RE type (typically by NP category). For each act of referent tracking, the referents were then were coded for accessibility using a system drawing on work by Toole (1996). This coding system provides a measure of referent accessibility that is independent of linguistic form, thereby avoiding the risk of circular reasoning (see discussion in H.-Y. Kim, 2000; Ryan, 2012; Tomlin, 1990). Toole applied the original system to widely varying types of data, with the findings indicating that "[t]he factors which affect referential choice are universal and apply regardless of genre" (1996, pp. 285-286)ⁱⁱ.

For reasons of space, the refined system is only briefly sketched here (see Ryan, 2012, 2015 for detailed explanation). Each referent is scored as the sum total of eight weighted criteria which either enhance or reduce accessibility, producing a single aggregated number between -1 and 8; since both -1 and 8 are rare, these are conflated with the scores at 0 and 7 respectively. These are reported as *degrees of accessibility*, ranging from D0 to D7. The assumption is that the scores roughly correspond to an interactants' sense of referent accessibility, in the sense that the higher the number, the higher the presumed accessibility of the referent.

The first and most heavily weighted criterion combines the concepts of distance (between an anaphor and its antecedent) and unity. Weighting for distance is given to referents mentioned in the current or previous proposition. Weighting for unity is given for referents mentioned in the current or previous 'episode': in film retelling tasks, as reported in Ryan (2015), these are demarcated by the clear temporal boundaries that exist between scenes; for free-flowing interview and conversation data, this concept has been reframed loosely as *narrative or thematic episode*, with episodes identified as starting with the signalling and establishment of major new topics and ending with their closure. In the present data, in many cases these were marked by disjunctive markers (e.g. "Oh, another thing I wanna tell you"), news announcements, or questions following silences (see Wong & Waring, 2010, for a summary of such practices). Topics in the interviews typically lasted at least 2-3 pages of transcript data, and sometimes considerably more.

Also requiring some further specification was the concept of discourse topicality (labelled global topicality in Ryan, 2015). Discourse topics were defined as the person most central to the thematic episode. In the majority of cases, each episode involving person reference had an obvious (and single) person who was readily identifiable as the main topic.

Though the system is not without limitations, its application to NS data reveals distributions of high- and low-accessibility markers in general accordance with Accessibility Theory (Ryan, 2015). Drawing on Ryan (2012, 2015), Table 1 below presents an overview of the RE types associated with each of the accessibility codes in L1 English, alongside the frequency with which other RE types are used. As displayed, accessibility contexts D5 to D7 represent high degrees of accessibility, where NS usually opt for pronouns or ø, while in contexts D0 to D2, they overwhelmingly avoided such forms, opting instead for names and determiner + noun combinations (with or without further modification). The intermediate range contexts of D3 and D4 appear to lie near a juncture allowing greater variation, though with LAMs clearly preferred at D3 and a slight tendency for HAMs at D4.

Table 1: Accessibility codin	ng in English	(based on Ryan,	2012, 2015)
	0 0	(·····,	- , ,

Accessibility range	Code(s)	Associated REs	Alternative RE options (Ryan, 2015)
High	D7	pn, ø	<1%
	D6	pn, ø	7%
	D5	pn, ø	12%
Intermediate	D4	pn, LAM	pn = 56%, LAM = 44%
	D3	LAM, pn	LAM = 74%, pn = 26%
Low	D2	LAM	6%
	D1	LAM	4%
	D0	LAM	0%

 ϕ = zero anaphora, pn = pronoun, LAM = low-accessibility marker (names, determiner + noun)

Findings: Longitudinal data

RE types used

Table 2 presents an overview of the RE types that Yoona used for referent tracking across the three interviews.

	ø	he/she	him/her	Name	<i>the/my/her</i> + NP	Bare NP	Other	Total
2016	5	121	27	19	5	2	3	181
2017	8	126	23	1	15	3	3	168
2018	0	100	27	7	7	0	3	144

Table 2: RE types used in referent tracking

Most notable in these data is the high proportion of pronouns used in each of the three interviews. Subject pronouns, object pronouns and zero anaphora (ϕ) together accounted for more than 80% of all references, with the proportion climbing slightly over the period (2016 = 81.3%; 2017 = 88.7%; 2018 = 88.2%). As explored further below, this suggests there was little if any over-explicitness, but perhaps substantial under-explicitness. In contrast to pronouns, ϕ was noticeably infrequent across all three interviews, with no cases at all in 2018. This is rather surprising given both their legitimate use in English (mainly in coordinate constructions with co-referential subjects), and their wide distribution in Korean discourse. The differing frequencies in the use of names can be accounted for by the nature of the stories told, and especially the number of referents known to the interviewer by name.

Across all three interviews, there were very few errors of form in the production of REs. There were, for instance, no pronoun errors (e.g. production of *he* instead of *she*) and few referential uses of bare nouns (e.g. *tutor* in place of *the tutor*) with none at all in 2018. In relation to the first research question, then, overall these initial figures offer little indication of changes in Yoona's referring behaviour over the $2\frac{1}{2}$ years.

Accessibility marking

It seemed clear from all three interviews that Yoona was skilled in tracking references, with relatively few indications of miscommunication or obvious infelicities. Here, this impression is explored by analysing the distribution of REs by accessibility context (D0-D7). Table 3

presents this overall distribution, with percentages provided for high-accessibility markers (HAMs) (subject, object, possessive pronouns and ø).

	2016		20	2017		018
	Total	% HAMs	Total	% HAMs	Total	% HAMs
D7	24	100%	48	100%	9	100%
D6	45	98%	52	100%	38	100%
D5	46	100%	16	100%	45	100%
D4	20	95%	16	100%	8	75%
D3	6	83%	12	83%	9	89%
D2	6	17%	6	67%	5	60%
D1	9	44%	6	33%	3	33%
D0	10	0%	4	0%	3	0%

Table 3: Distribution of pronouns by accessibility context (Yoona)

These figures lead to five rather striking observations:

- (1) Yoona's use of pronouns within the higher accessibility contexts was highly consistent.
- (2) Pronoun use dominated in all but the lowest accessibility contexts.
- (3) There is effectively no evidence of over-explicitness in the data.
- (4) There is substantial evidence of under-explicitness.
- (5) There is strong evidence that Yoona's patterns of accessibility marking remained highly stable across the 2 ¹/₂ years of the study.

Elaborating firstly on the first three observations, across the three interviews Yoona made 323 references in contexts D5-D7 and all but one of these was with a high-accessibility marker, nearly always pronoun. The only exception was a single use of a name at D6 in the first interview. In 2017, such consistency was maintained across an even wider distribution of contexts, encompassing all 132 references between D4-D7. This contrasts strikingly with

previously reported distributions both in L1 English (see Table 1 above) and L2 English. While intermediate and advanced L2 English is routinely over-explicit, this is clearly not the case with Yoona. Furthermore, Yoona's consistency in selecting pronouns appears markedly greater than has been reported for L1 English; native English speakers are highly consistent in using HAMs only in context D7, with full NPs accounting for about one in ten references at D5-D6 (Ryan, 2015; see also Table 1 above). This can be partly explained by NSs occasionally varying their RE selection for purposes beyond ensuring that referents are identifiable, for instance to achieve particular stylistic effectsⁱⁱⁱ, to signal the structuring of discourse (Vonk, Hustinx, & Simons, 1992), or to provide additional information such as stance, as in *poor old Gladys* (Stivers, 2007). In not applying such strategies, Yoona's referring practices at D4-D7 are best characterised as being consistently *minimally explicit*, as they are felicitous but minimally informative^{iv}.

In the lower-accessibility contexts of D1 to D3, it appears that Yoona was frequently *under-explicit*, selecting pronouns where full NPs are felicitous. At D3, where full NPs might be expected to account for around 75% of references, her selections were dominated by pronouns (23/27; 85%). Under-explicitness is particularly unequivocal at D2 and D1, where almost half of her total references were by pronouns (15/35); by comparison, they are relatively rare in NS speech and attributable to "occasional misjudgements during unplanned speech" (Ryan, 2015, p. 845).

The consistency in Yoona's RE selections confirm that her minimal and under-explicitness are not the result of random infelicities, but instead result from highly rule-governed behaviour in which she systematically marked less-accessible referents as though they were more highly-accessible. Thus, Yoona's system of accessibility marking appears to involve a slightly different mapping of NP-type to degree of accessibility, with pronouns being used for a wider range of accessibility. This appears particularly clear within the intermediate range of D3-D4, which marks a juncture in L1 English between the use of high- and low-accessibility markers, but where Yoona overwhelmingly used pronouns across the three years. More tentatively (due to the small data set) but also more intriguingly, in 2017 and 2018 this is also the case at D2, where pronouns appear more unequivocally infelicitous and are perhaps prone to triggering miscommunication (as will be discussed). The fact that such pronoun use

became more frequent after 2016 could suggest a slight shift to even greater underexplicitness over time.

Yoona's system of accessibility marking can therefore be described for each of the three data collection points of the study, though more tentatively for the lower accessibility contexts where there is less data. This is presented in Table 4 below alongside expected felicitous RE selection based on distributions in NS data (based on Ryan, 2015).

Context	Felicitous REs		Yoona	
	-	2016	2017	2018
D7	pn, ø	pn	Pn	pn
D6	pn, ø	pn	Pn	pn
D5	pn, ø	pn	Pn	pn
D4	pn, LAM	pn	Pn	pn
D3	LAM, pn	pn	Pn	pn
D2	LAM	LAM	Pn	pn
D1	LAM	LAM	LAM	LAM
D0	LAM	LAM	LAM	LAM

Table 4: Yoona's system of accessibility marking

In terms of longitudinal development, there is thus the tentative suggestion that Yoona's system of accessibility marking may have become more under-explicit after 2016 in terms of using pronouns in context D2. Overall, however, the most notable observation is again the overall stability of the system over the 2 ¹/₂ years.

It is also worth mentioning that, like native speakers in previous studies, referent accessibility was a very strong predictor of Yoona's RE selection, with no additional effect detected within the contexts of topic shift or focus position. By contrast, language learners are elsewhere reported as tending to be particularly over-explicit in both of these contexts (e.g. Ryan, 2015).

From these analyses, a picture emerges in which Yoona's referring behaviour differs markedly from what was anticipated based on the previous literature. Over-explicitness, which is so characteristic within the speech of many SLLs, seems entirely absent from Yoona's talk except perhaps in the limited sense of under-using ø. Instead, her talk is characterised by a notable tendency towards minimal informativity by way of extensive use of pronouns, at times erring on the side of under-explicitness. To confirm this finding, the following subsection provides qualitative analyses of extracts coded as containing underexplicit pronouns.

Qualitative analyses

As will be demonstrated, the qualitative analyses support the evidence for an absence of overexplicitness and relatively frequent under-explicitness. To convincingly illustrate the latter, in most cases somewhat extended extracts of text are required to give a sense of distance, competition, unity and so on. These analyses also suggest some longitudinal developments that are not captured in the quantitative analyses.

As discussed in the methodology section, plural references, references in reported speech and references made by the interviewer were not coded for accessibility, but are taken into account when determining the accessibility of subsequent references. For ease of reading, Yoona's references are bolded. The identity of the referents are distinguished in subscript (e.g. _A or _B) and the accessibility scores are presented in superscript.

Examples of under-explicitness are readily identifiable across all three interviews. Extract 1 begins 31 minutes into the final interview, and illustrates Yoona's tracking of multiple referents with minimal, and perhaps occasionally under-explicit reference. The extract

contains references to four individuals within a short space of time. Included are a pronominal reference to _(B)Grace coded at D2 (line 10), which was successful despite potential ambiguity from referent _(A) in line 08. Similarly, the pronoun in line 11 (coded D3) initially seems ambiguous given the competition from _(B) and _(D) but soon becomes clear once _(B)Grace is ruled out by the non-co-referential "before _(B)Grace". The final pronoun in line 17 is also minimally explicit (coded D3) but is clear within the communicative context Yoona describes.

Extract 1

01	Y:	actually, um, (B)Grace(intro) was in semester two with me
02		but I had hadn't talked to $_{(B)}her^5$ until semester six,
03	J:	oh
04	Y:	because we were like in different stream, and
05	J:	oh yep, yep
06	Y:	I didn't really catch up with anybody, so
07	J:	yep
08	Y:	except (A)my Filipino friend ⁻¹ and yeah,
09		and I yeah, I felt international students are all the same,
10		but but (B) <u>her² (C)friend^(intro)</u> was quite close with (D) <u>Emma⁻¹</u> ,
11		so I got close to (C) <u>her</u> ³ , before (B) <u>Grace</u> ³ ,
12		and $_{(C)}$ <u>she³</u> told me that – before going to placement we met at the library –
13		and then (C)she ⁶ said 'you going to placement with (B)Grace, together like?,'
14		'oh, do I?' and then 'How is (B)she? I'm so worried about my English,
15		like I'm gonna met my supervisors
16		and then like I have to become independent,'
17		and then (C)she's ³ like 'you're fine',

Consideration of such extracts supports the conclusion that Yoona frequently used pronouns in contexts where a fuller form might be expected. However, despite differing from L1 norms, Yoona's management of multiple referents in such an economical and yet communicatively successful way appears skilful.

Extract 2 is from 13 minutes into the interview, and illustrates the ambiguity that occasionally arises from under-explicitness. Immediately prior to this, Yoona had been discussing the impact of her closest friend (A) needing to repeat a semester and therefore being placed in different classes. She then speaks of the consequent need to communicate more with others. In line 14, "another kiwi girl" (B) is introduced into the discourse. In line 16, she states that "she's actually like a Filipino kiwi". Here, *she* seems felicitous for referring to (B), who had accessibility D5; it seems under-explicit for (A), who was coded at D3. This caused some confusion as I knew that (A) was ethnically Filipino, hence the clarification request initiated in line 19 to determine whether they were both Filipino. Further ambiguity arose from another under-explicit (D3) pronoun in line 28.

Extract 2: October 2018

01	Y:	but since (Y & A)we split from each other, I had to make a friend?,
02	J:	yeah
03	Y:	and I had to talk to other people, or other girls,
04		other kiwi students, yeah.
05	J:	mm
06	Y:	yeah y'kn[ow, I think
07	J:	[so that, that was the turning point,
08		so[rt of when you were forced to kinda
09	Y:	[yeah I think,
10		and, yeah, turning point to become a little bit more independent,
11	J:	yeah,
12	Y:	yeah,
13		but since, yeah, that semester, yeah, semester five,
14		I was with $_{(B)}$ another kiwi girl, and sh- I learned a lot from $_{(B)}$ her ⁵ ,
15		the-like the attitudes and the communication way, and skill,
16		I learned a lot, and, $_{(A)}$ <u>she's a – she</u> 's ³ actually like a Filipino kiwi,
17		but (A) <u>she</u> 's ⁶ like – they have like different communication way,
18		so, yeah, so
19	J:	oh, just to clarify, you mean the- they're(A & B) both Filipino kiwi?
20	Y:	no no

- Ryan, J. (2020). Under-explicit and minimally explicit reference: Evidence from a longitudinal case study. In J.
 Ryan & P. Crosthwaite (Eds.) *Referring in a second language: Studies on reference to person in a multilingual world* (pp. 100-118). London: Routledge.
- 21 J: nah nah (A)<u>your original friend</u> was and (B)<u>the other</u> is

22	Y:	kiwi,
23	J:	just pakeha-kiwi?
24	Y:	yeh
25	J:	yup
26	Y:	yeh,
27	J:	mm
28	Y:	but $_{(A)}$ she ³ has grown up here so, $_{(A)}$ she ⁽⁶⁾ only speak English too, so
29	J:	yeah
30	Y:	but it was kinda different experience,

Extract 3 below illustrates minimal – and perhaps under-explicit – reference in a somewhat different context. At about 11 minutes into the interview, the topic switched to recalling where our discussions had left off in the last interview two months previously, and specifically a dilemma Yoona had described regarding her working schedule: she had wanted to keep working under the same supervisor, but this would have meant changing to a night shift, creating other difficulties. In this extract, the focus of interest is the introduction of this supervisor in line 09. Of interest is that Yoona uses a pronoun to do so, and that – perhaps counter-intuitively – this proves communicatively successful. As a referent introduction, no accessibility code is given.

Extract 3: October, 2018

01	J:	yeh, you- I remember you changed your time didn't you,
02		because you were (0.6) working nights, and then
03		(1.0)
04		or ev[enings and then changed to
05	A:	[oh yeah
06		(0.7)
07		ye[ah
08	J:	[or day time, yu[p
09	A:	[yeh, (0.3) and I text (A)her,
10		(0.6)

- 11 J: mm
- 12 A: tell (A)her that I preferred PM shift, to do my assignment,
- 13 and then (A)she #said 'oh that's totally fine,'
- 14 J: yeh

As suggested above, this reference seems rather curious as such *recognitional* introductions are associated with the use of full NPs and sometimes lengthy introduction sequences (Ryan, 2016; Smith, Noda, Andrews, & Jucker, 2005). In this case, the previous reference to the supervisor was 58 days earlier. Yet, as the interlocutor, there was little or no processing strain in interpreting this pronoun. In accounting for this, the most relevant consideration is referent accessibility rather than discourse status (as introduced, maintained or re-introduced). Specifically, mutual ground had been collaboratively established over the situation being discussed, with a demonstration of hearer-understanding through lines 01-04. In the establishment of a suitably specific context, the accessibility of the supervisor was increased. Yoona thus correctly judged that a pronoun would be sufficiently clear.

To briefly summarise, then, the qualitative analysis supports and strengthens the quantitative findings regarding extensive minimal and – at times – under-explicit reference. Such evidence of under-explicitness occurs across all three interviews^v. However, my overall impression was that cases of under-explicitness from the 2017 and 2018 interviews tended to be more striking and perhaps more obviously infelicitous than those in 2016. This forced a closer inspection of the contexts in which these occurred. The most notable difference appeared to be that all of the anecdotes recounted in the 2016 interview required the concurrent tracking of just one or two referents; these could therefore be readily interpreted with little risk of miscommunication. By contrast, in both the 2017 and 2018 interviews there were a number of stories involving a concurrent focus on three or even more referents (see, for example, Extract 1). Since these engender a greater risk of miscommunication from ambiguous RE selection, instances of under-explicitness often appeared more conspicuous by way of (potential) vagueness and therefore requiring additional processing effort for the hearer. This could simply reflect the nature of the stories that Yoona had to tell in 2016; however, it also seems likely that in the later interviews she *chose* to tell more complex

stories, or to tell them in richer and more referentially complex ways. This seems very plausible given her greater communicative experience and growing interactional competence.

There were some other subtle indications of developments over time. For instance, Extract 4 below from 2016 illustrates a feature only observed in this first interview, whereby a vague plural *they* becomes the antecedent of a singular reference (*she*); this was observed three times in the first interview but not at all in later interviews. Prior to this extract, Yoona had vaguely mentioned everyone, which was inferable as meaning her classmates; this was shortly followed by a co-referential *they*. Line 01 below occurred ten turns (22 propositions) later, and includes a further plural use of *they*. This then becomes the antecedent of a further they in line 02; however, here the plural pronoun evidently relates to a singular referent, who has a daughter "crazy about K-pop" and wanting "to go to Korea". In line 07, this is followed by she (line 7). The interpretation of she requires inferring that it is one of the classmates, and specifically the individual (*they*) who spoke of a daughter and K-pop. This transition from a vague plural to a specific singular reference seems curious and is somewhat different to the various subclasses of antecedentless pronouns discussed elsewhere (e.g. Gundel, Hedberg, & Zacharski, 2005; Yule, 1982). Nevertheless, it presented no apparent communicative difficulty, and may in fact represent an effective and economical means of temporarily introducing peripheral hearer-new individuals into discourse.

Extract 4: July 2016

01	A:	and (B*)they tried to make me like feel comfortable
02		and $_{(B^*)}$ they [said] 'oh $_{(A)}$ my daughter really really crazy about K-pop
03		so (A)she really want to go to Korea or something'
04	J:	oh
05	A:	so I said 'oh really' and $\underline{0}$ [we] talked to each other,
06	J:	uh-huh
07	A:	and $_{(B)}$ she ⁽⁶⁾ was really nice, and actually they really helped me yesterday,
08		cause I wasn't here Monday so I don't – I have no idea about $_{(C)}\underline{her}^{(1)}$ lectures,
09		so I said I wasn't here

Discussion and conclusion

In relation to the two research questions, in the nearly 2 ½ year span of the study there is less evidence than expected of (1) shifts in Yoona's system of RE use and (2) patterns of accessibility marking. The development of Yoona's RE system included the apparent elimination of ø and bare nouns by 2018, and by 2017 dropping the practice of occasionally using singular anaphors (*she*) to specify one of from a set of vague plural antecedents (*they*). A more substantial development was that by 2017, Yoona was engaging in more complex concurrent tracking of multiple referents, which may reflect greater overall confidence and experience in speaking English. In terms of accessibility marking, there is tentative evidence of increasing use of pronouns in the low-accessibility context D2.

However, the most striking findings overall were Yoona's strong tendency towards minimal and under-explicitness and that she was very seldom if ever over-explicit. These were stable features of her accessibility marking across the length of the study, and are counter to the predictions arising from previous studies. While previous evidence of under-explicitness has mainly involved infelicitous use of ϕ (in place of pronouns), here it involved the use of pronouns where names and *the* + noun are felicitous, despite the use of pronouns being very infrequent in Korean.

The stability of Yoona's referent tracking system is of interest given the 2 ½ span of the study, and her extensive use of and exposure to English during this period^{vi}. Numerous factors have been proposed to account for such long-term stability in SLL language systems (see Long, 2003 for an overview) but apparently little specifically in relation to L2 pragmatics. In this case it seems helpful to consider the competing interactional demands of achieving referential clarity (or *recognition*) and being economical (Levinson, 2007; Sacks & Schegloff, 2007). Achieving clarity can require providing substantial descriptive information, while economy promotes brevity. From this perspective, Chini's (2005) early pragmatic and lexical stage is characterised by lapses in clarity through over-use of ø and omission of determiners. Among factors driving further development will be clarification requests and interactional repairs, as well as positive and negative evidence from language exposure. At the following (over-) explicit lexical stage, clarity is emphasised but at a cost to economy,

with over-explicit references and other redundant information slowing communication. This may eventually be a source of frustration as it exasperates the pre-existing communicative 'bottleneck' in which the mind processes language considerably faster than the ability to verbalize it (Levinson, 2000). Subsequent development may therefore be largely driven by internal and external pressures for economical communication, alongside positive evidence from language exposure.

In Yoona's case, her strategy of being highly economical appears to have been largely successful, despite at times being infelicitous by NS standards. There is no contradiction here: as argued by Kasper (1997, pp. 355-356), it is simply flawed reasoning to assume L2 speech acts are communicatively problematic simply because they differ from L1 norms. Here, it appears that Yoona successfully monitored her RE selections for clarity and was well-attuned to the interlocutor's ability to recover the referent. Even notably under-explicit REs (e.g. those at D1 and D2) seldom proved communicatively problematic, though some may have required more processing effort from the interlocutor. Indeed, among the 493 references examined, there were only five clarification requests (1.0%). Since these also occur relatively frequently in NS interactions, this can be presumed to be fairly target-like and likely too few to prompt a behavioural change. Conversely, the 488 references that passed unremarked (99.0%) provide Yoona with evidence of success, while also facilitating the economical flow of discourse. In short, then, the stability of Yoona's accessibility marking is likely a reflection of its overall success in terms of both relative clarity and economy.

In further accounting for the differences between these and previous findings, the limitations of the study need to be acknowledged along with reflections on how the data were generated. The most important limitations relate to the limited number of references examined. While the interviews probably provide a sufficient number overall (just under 500), as with previous studies, these are concentrated particularly among the higher accessibility contexts (D5-D7), with references to less-accessible individuals remaining greatly under-researched.

Particularly relevant features of the interview data include the familiarity between the interactants, the relaxed setting, the participant's control over the stories she chose to tell,

extensive opportunities to rehearse her ideas prior to the interview, the more naturalistic setting, and the communicative focus being co-constructed by the interactants rather than predetermined by the task. Each of these features likely serves to reduce communicative pressure. Such unpressured performance allows participants "the opportunity to conceptualise, formulate and articulate their messages with some care" (Ellis, 2005, p. 165), potentially allowing greater precision in RE selection.

By contrast, the vast majority of related studies have made use of elicitation tasks, such as retellings of silent films and picture sequences conducted in classroom or laboratory-type settings (e.g. Ahrenholz, 2005; Chini, 2005; Gullberg, 2006; Hendriks, 2003; Ryan, 2015). Such tasks lend themselves to having pre-determined and comparatively inflexible criteria for accuracy in meaning, against which participants feel they are measured. Thus, they involve pressured performance, which has been shown to have a negative impact on language complexity and accuracy (e.g. Ellis & Yuan, 2005). Since this arises from the additional demands required at the level of planning and organising information, it would not be surprising if this also had a negative effect on felicitous RE selection. It could be, then, that over-explicitness is more common in traditional elicitation tasks than in naturally occurring L2 speech. Indeed, even without the challenge of second language use, both planning load and memory load are also associated with greater explicitness in L1 English (Arnold, 2010). With this in mind, further studies of L2 reference are recommended using data generated in unpressured performance, such as through informal interviews, discussions and conversation.

References

- Ahrenholz, B. (2005). Reference to persons and objects in the function of subject in learner varieties. In H. Hendriks (Ed.), *The structure of learner varieties* (pp. 19-64). Berlin, Germany: Mouton de Gruyter.
- Ariel, M. (1990). Accessing noun phrase antecedents. London, England: Routledge.
- Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schiperoord, & W.
 Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 29-87). Amsterdam, The Netherlands: John Benjamins.

- Ryan, J. (2020). Under-explicit and minimally explicit reference: Evidence from a longitudinal case study. In J.
 Ryan & P. Crosthwaite (Eds.) *Referring in a second language: Studies on reference to person in a multilingual world* (pp. 100-118). London: Routledge.
- Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, 4(4), 187-203. doi:10.1111/j.1749-818X.2010.00193.x
- Bach, K. (2008). On referring and not referring. In J. K. Gundel & N. Hedberg (Eds.), *Reference: Interdisciplinary perspectives* (pp. 13-58). Oxford, UK: Oxford University Press.
- Broeder, P. (1991). *Talking about people: A multiple case study on adult language acquisition*. Amsterdam, The Netherlands: Swets & Zeitlinger.
- Chini, M. (2005). Reference to person in learner discourse. In H. Hendriks (Ed.), *The structure of learner varieties* (pp. 65-110). Berlin, Germany: Walter de Gruyter.
- Crosthwaite, P. (2013). An error analysis of L2 English discourse reference through learner corpora analysis. *Linguistic Research*, *30*(2), 163-193.
- Crosthwaite, P. (2014). Differences between the coherence of Mandarin and & Korean L2 English learner production and English native speakers: An empirical study.
 (Doctoral doctoral thesis), University of Cambridge, UK.
- Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3-34). Amsterdam, Netherlands: John Benjamins.
- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167-192). Amsterdam, The Netherlands: John Benjamins.
- Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, *56*(1), 155-196. doi:10.1111/j.0023-8333.2006.00344.x
- Gundel, J. K., Hedberg, N., & Zacharski, R. (2005). Pronouns without NP antecedents: How do we know when a pronoun is referential? In A. Branco, T. McEnery, & R. Mitkov (Eds.), *Anaphora processing: Linguistic, cognitive and computational modelling* (pp. 351-364). Amsterdam, The Netherlands: John Benjamins.
- Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 291-326). Berlin, Germany: Mouton de Gruyter.
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition, 24*(3), 387-418. doi:10.1017/S0272263102003029

- Ryan, J. (2020). Under-explicit and minimally explicit reference: Evidence from a longitudinal case study. In J.
 Ryan & P. Crosthwaite (Eds.) *Referring in a second language: Studies on reference to person in a multilingual world* (pp. 100-118). London: Routledge.
- Kasper, G. (1997). Beyond reference. In G. Kasper & E. Kellerman (Eds.), Communication strategies: Psycholinguistic and sociolinguistic perspectives (pp. 345-360). London, UK: Longman.
- Kim, H.-Y. (2000). *Acquisition of English nominal reference by Korean speakers*. (Doctoral dissertation), University of Hawai'i at Manoa, HI.
- Kim, S.-H. (2013). A corpus-based analysis of discourse anaphora in English and Korean: A neo-Gricean pragmatic approach. (PhD Dissertation), The University of Texas at Arlington,
- Klein, W., & Perdue, C. (1992). *Utterance structure: Developing grammars again*. Amsterdam, The Netherlands: John Benjamins.
- Knowler, R. (2010, July 29). All Blacks staff spell it out for Isaac Ross. *The Press*. Retrieved from http://www.stuff.co.nz/sport/rugby/provincial/3967950/All-Blacks-staff-spell-it-out-for-Isaac-Ross
- Leclercq, P., & Lenart, E. (2013). Discourse cohesion and accessibility of referents in oral narratives: A comparison of L1 and L2 acquisition of French and English. *Discours: Revue de Linguistique, Psycholinguistique et Informatique, 12*, 3-31. doi:10.4000/discours.8801
- Levinson, S. C. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Levinson, S. C. (2007). Optimizing person reference perspectives from usage on Rossel Island. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 29-72). Cambridge, UK: Cambridge University Press.
- Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Malden, MA: Blackwell.
- Lozano, C. (2009). Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus. In N. Snape, Y.-k. I. Leung, & M. Sharwood Smith (Eds.), *Representational deficits in SLA: Studies in honor of Roger Hawkins* (pp. 127-166). Amsterdam, The Netherlands: John Benjamins.
- Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In M.

Alonso-Ramos (Ed.), *Spanish learner corpus research: Current trends and future perspectives* (pp. 235-265). Amsterdam, The Netherlands: John Benjamins.

- Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: Pronominal subjects in Greek learners of Spanish. *Journal of Psycholinguistic Research*, 47(2), 411-430. doi:10.1007/s10936-017-9541-8
- Lumley, J. R. (2013). Pragmatic perspectives on the second language acquisition of person reference in Japanese: A longitudinal study. (PhD Doctoral thesis), Newcastle University, Newcastle, UK.
- Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge, UK: Cambridge University Press.
- Mitkov, R. (2002). Anaphora resolution. London, UK: Longman.
- Nakahama, Y. (2003). Development of referent management in L2 Japanese: A film retelling task. *Studies in Language and Culture*, 25(1), 127-146. Retrieved from http://www.lang.nagoya-u.ac.jp/proj/genbunronshu/125-121/nakahama.pdf.
- Nakahama, Y. (2009). Cross-linguistic influence on referent introduction and tracking in Japanese as a second language. *The Modern Language Journal*, 93(2), 241-260. doi:10.1111/j.1540-4781.2009.00859.x
- Nakahama, Y. (2011). Referent markings in L2 narratives: Effects of task complexity, learners' L1 and proficiency level. Tokyo, Japan: Hituzi Syobo.
- Odlin, T. (2003). Cross-linguistic influence. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 436-486). Oxford, England: Blackwell.
- Ryan, J. (2012). Acts of reference and the miscommunication of referents by first and second language speakers of English. (PhD doctoral thesis), University of Waikato, Hamilton, New Zealand.
- Ryan, J. (2015). Overexplicit referent tracking in L2 English: Strategy, avoidance, or myth? Language Learning, 65(4), 824-859. doi:10.1111/lang.12139
- Ryan, J. (2016). Introducing referents for recognition: L2 pragmatic competence and miscommunication. *Journal of Pragmatics*, 97, 55-73. doi:10.1016/j.pragma.2016.04.005
- Sacks, H., & Schegloff, E. A. (2007). Two preferences in the organization of reference to persons in conversation and their interaction. In N. J. Enfield & T. Stivers (Eds.),

Person reference in interaction: Linguistic, cultural, and social perspectives (pp. 23-28). Cambridge, UK: Cambridge University Press.

- Smith, S. W., Noda, H. P., Andrews, S., & Jucker, A. H. (2005). Setting the stage: How speakers prepare listeners for the introduction of referents in dialogues and monologues. *Journal of Pragmatics*, 37(11), 1865-1895. doi:10.1016/j.pragma.2005.02.016
- Song, J. J. (2005). *The Korean language: Structure, use and context*. Abingdon, UK: Routledge.
- Stivers, T. (2007). Alternative recognitionals in person reference. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 73-96). Cambridge, UK: Cambridge University Press.
- Takeuchi, M. (2014). Subject referential expressions and encoding of referential status in l2 narrative discourse by L1-English learners of Japanese. (PhD Dissertation), Indiana University, Bloomington, IN.
- Tomlin, R. S. (1990). Functionalism in second language acquisition. *Studies in Second Language Acquisition*, *12*(2), 155-177. doi:10.1017/S0272263100009062
- Toole, J. (1996). The effect of genre on referential choice. In J. K. Gundel & T. Fretheim (Eds.), *Reference and referent accessibility* (pp. 263-290). Amsterdam, The Netherlands: John Benjamins.
- Vonk, W., Hustinx, L., & Simons, W. H. G. (1992). The use of referential expressions in structuring discourse. In J. Oakhill & A. Garnham (Eds.), *Discourse representation* and text processing: Special issue of 'Language and Cognitive Processes' (pp. 301-333). Hove, UK: Lawrence Erlbaum.
- Wong, J., & Waring, H. Z. (2010). Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers. New York, NY: Routledge.
- Yule, G. (1982). Interpreting anaphora without identifying reference. *Journal of Semantics*, 1(4), 315-322. doi:10.1093/jos/1.3-4.315

ⁱ For greater detail on aspects of Korean reference, the interested reader is directed to H.-Y. Kim (2000) and S.-H. Kim (2013).

ⁱⁱ Parallel systems are found for anaphora resolution, originating within the fields of computational linguistics and natural language processing (see for example Mitkov, 2002). Such systems are designed for automated interpretation rather than accounting for speaker behaviour.

ⁱⁱⁱ This is illustrated in the following media interview with a sports coach: "<u>Isaac</u> knows clearly what <u>he</u> has got to do. I think everyone that has watched <u>Isaac</u> play knows what <u>Isaac</u> has got to do. And <u>Isaac</u> really needs to understand that the only person that can fix <u>Isaac</u>'s problems is <u>Isaac</u>" (Knowler, 2010, July 29). ^{iv} Except in the sense that pronouns are used where ø would be appropriate.

Except in the sense that pronouns are used where \emptyset would be appropria

^v Note for instance below in Extract 4 (2016) the use of 'her' in line 08.

^{vi} Although there is insufficient space for details here, as expected there is clear evidence in other domains of her gradually increasing linguistic and pragmatic competence over this period, including her fluency, vocabulary range, grammatical range and morphological accuracy.